Generalized Trust in the Mirror: An Agent-Based Model on the Dynamics of Trust

Klein, Dominik; Marx, Johannes

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Historical Social Research
Historische Sozialforschung

Dominik Klein & Johannes Marx:

Generalized Trust in the Mirror. An Agent-Based Model on the Dynamics of Trust.

doi: 10.12759/hsr.43.2018.1.234-258

Published in:

Historical Social Research 43 (2018) 1

Cite as:

Klein, Dominik, and Johannes Marx. 2018. Generalized Trust in the Mirror. An Agent-Based Model on the Dynamics of Trust. *Historical Social Research* 43 (1): 234-58. doi: 10.12759/hsr.43.2018.1.234-258.

Historical Social Research
Historische Sozialforschung

All articles published in HSR Special Issue 43 (2018) 1:
Agent-Based Modeling in Social Science, History, and Philosophy.

Generalized Trust in the Mirror. An Agent-Based Model on the Dynamics of Trust

Dominik Klein & Johannes Marx [*]

Abstract: *»Gesellschaftliches Vertrauen in Reflektion. Eine agentenbasierte Modellierung von Vertrauensdynamiken.«* High levels of trust have been linked to a variety of benefits including the well-functioning of markets and political institutions or the ability of societies to solve public goods problems endogenously. While there is extensive literature on the macro-level determinants of trust, the micro-level processes underlying the emergence and stability of trust are not yet sufficiently understood. We address this lacuna by means of a computer model. In this paper, conditions under which trust is likely to emerge and be sustained are identified. We focus our analysis mainly on the individual characteristics of agents: their social or geographical mobility, their attitude towards others or their general uncertainty about the environment. Contrary to predictions from previous literature, we show that immobile agents are detrimental to both, the emergence and robustness of trust. Additionally, we identify a hidden link between trusting others and being trustworthy.

Keywords: Generalized trust, agent-based modeling, social simulation, trust-game.

1. Introduction

Overcoming collective action problems is a core challenge for large and diverse societies. A prominent debate, initiated by Ostrom's seminal work (1990) focuses on the question of when and under which condition collective action problems can be solved endogenously, through self-coordination. In this strand of literature, trust has been identified as a key determinant for enabling societies to solve such problems autonomously and efficiently. Moreover, high levels of trust have been linked to a variety of social and individual benefits including the performance of political institutions (Putnam 2000; Putnam, Leonardi and Nanetti 1994), economic capabilities of states (Knack and Keefer 1997), or health and a better quality of life (Hyyppä 2010). In light of these

[*] Dominik Klein, Political Science Department, University of Bamberg, Feldkirchenstrasse 21, 96052 Bamberg, Germany; dominik.klein@uni-bamberg.de.
Johannes Marx, Political Science Department, University of Bamberg, Feldkirchenstrasse 21, 96052 Bamberg, Germany; johannes.marx@uni-bamberg.de.

Historical Social Research 43 (2018) 1, 234-258 | published by GESIS
DOI: 10.12759/hsr.43.2018.1.234-258

effects, it comes as no surprise that we find a growing interest in the determinants of trust over recent years. Much effort has been expended identifying a range of factors relevant for the emergence and stability of trust. These include institutional factors, but also a variety of cultural, societal and individual variables (Hooghe and Stolle 2003; Kornai, Rothstein and Rose-Ackerman 2004). We complement these external determinants with a closer look at the endogenous dynamics of trust or distrust. Far from being a stable phenomenon, trust is created through an ongoing complex dynamic process. For once, this process is interactive and self-reinforcing: trust creates trust. But how exactly do such iterated interactions on the micro-level combine to societal trust? And which role do macroscopic factors such as agents' mobility or their shared cultural heritage play in this process? These are the questions we address in this paper. That is, we are not only interested in *which* factors contribute towards the emergence or destruction of trust, but also *how* they do so and how trust and distrust sustain.

We do so by developing an agent-based model in NetLogo (Wilensky 1999), building on current psychological models as well as insights from the social sciences. We see two major advantages in applying computer simulation to the emergence and dynamics of trust:

- First, a computer simulation helps to handle methodological problems inherent in empirical research on trust. What might seem tautological – trust reinforcing itself – and what might cause methodological problems in an empirical study can easily be disentangled with the help of computer simulations, which can also determine at what point in time an individual has acquired trust proper, and when her behavior merely reflects second or third level considerations.
- Second, we are interested in the quality of our theoretical knowledge of trust. We seek to understand the mechanisms that lead to low or high levels of trust in societies. We therefore build our simulations on well-established theories of the determinants of trust. Agents' mobility as well as the stability of their surroundings, for example, are held as central influence factors. But with our simulation we will demonstrate that empirical results about the direction of these factors' influence need not always be true. Low mobility and stable surroundings might sometimes increase trust, but our simulation suggests that this need not always be the case.

In what follows we will develop and analyze a dynamic model of social trust. We will first clarify our notion of trust in Section 2. In Section 3 we introduce our model. Section 4 contains an overview of the results, before we conclude in Section 5.

2. Theoretical and Empirical Research on Trust

Trust is a multifaceted concept. In some cases, it can describe the attitude among two complete strangers. It can also refer to the relationship between good friends. Some authors argue that trust can occur only in actions. Others hold that trust refers rather to a belief or attitude. Correspondingly, the current literature contains a multitude of theoretical approaches to trust (see for instance Sztompka 1999; Torche and Valenzuela 2011). The literature on social capital (Putnam et al. 1994; Uslaner 2002) distinguishes thick and thin notions of trust. The former refers to personalized attitudes and expectations towards well known, individual others. This thick notion of trust is grounded in a well-established social relation between the actors, based on acquaintance, joint past experience, institutional frames, or expectations of future interaction. The emergence of thick trust has been thoroughly studied with various theoretical and empirical models as well as computational simulations (Nooteboom, Klos and Jorna 2001). The thin notion of trust, conversely, refers to the general attitude towards strangers, anonymous and hitherto unknown members of society that we might not expect ever to see again. Faced with situations of thin trust, agents base their behavior on past experience in similar situations, demeanor, appearance or, more general, membership in certain social groups (Birk 2001). We are interested here particularly in this second notion of thin or *generalized* trust.

Other parts of the literature differ in their understanding of what it means to trust. One line of research holds that trust is a behavioral concept, that is or could be manifested in the actions of some agents. Other approaches argue that trust is a cognitive phenomenon, a belief or a family of beliefs. In this paper, we focus on a narrow definition of trust, following the latter alternative. We take trust as an expectation that may or may not be held by rational agents.[1] More specifically, trust is considered to be an agent's belief that a randomly chosen counterpart will act cooperatively in certain strategic situations. A high level of trust is essential for solving cooperation problems that arise in situations of strategic interdependence. The structure of such a situation can be characterized by the following conditions: A resource S is shifted from actor A, the trustor, to actor B, the trustee. The trustor's reason for this shift of resources is the hope or expectation to gain from that interaction. However, in shifting resources, actor A makes herself vulnerable. She will profit from this interaction only if her counterpart acts cooperatively.

[1] In our simulation, there is a close link between *trust* as a belief and *placing trust* as a behavior. Beliefs about the trustworthiness of some actor will be readily translated into some action or non-action towards that actor.

Trust games involve a crucial temporal asymmetry. The trustor shifts resources prior to learning about the trustee's response. The latter, in turn, does not need to decide on her behavior until the trustor has moved. If the trustee proves trustworthy, both parties receive a positive payoff. However, if the trustee turns out untrustworthy only the trustee benefits while the trustor ends up with the worst possible payoff. Thus, engaging in trust games is a conscious decision under risk. The trustor invests voluntarily and without guaranteed success. Furthermore, the trustee's benefit upon exploiting trust is higher than upon rewarding trust. That is, the trustee has no material incentive to act trustworthily. Of course, a rational trustor will engage in a trust game only if she expects the opponent to be trustworthy. But, given the temporal structure of the game, there is no guarantee that the trustee will act as expected. The trustor might assess her opponent's trustworthiness incorrectly, trusting some defector or refusing to play with a trustee who would have cooperated. In light of this structural uncertainty, a rational trustor will not have all-out beliefs about her opponent, judging him trustworthy or untrustworthy tout court. Rather, she will engage in some cautious, fine grained considerations, judging the trustee to be more or less trustworthy. To accommodate this complexity, we will define trust as a graded variable (Coleman 1994, 91-116), reflecting *how likely* the trustor judges her counterpart to be trustworthy. Within the framework of bounded rational choice theory, this situation can be represented as a trust game in extensive form with incomplete information (Buskens 2002). In this framework, the trustor's uncertainty about the trustee's behavior is represented as uncertainty about her payoff structure, as depicted in Figure 1.[2] In this model, uncertainty about the trustee's motivations is expressed as a draw by nature: With probability p, A's counterpart is not trustworthy, i.e., A will interact with a partner with a dominant strategy of defecting. $1 - p$ is the corresponding likelihood of playing with a trustworthy player, having cooperation as her dominant strategy.[3]

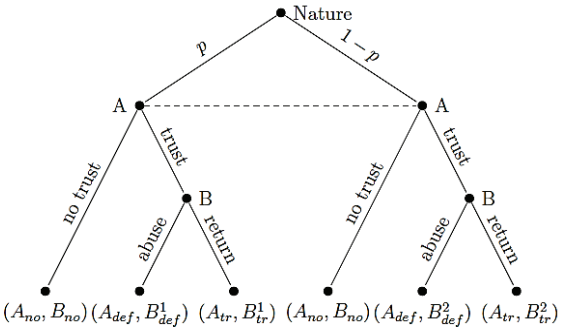
The central task for the trustor is thus to estimate the likelihood of being paired with a trustworthy partner. A trustor will agree to place trust in her counterpart if and only if she judges as high enough the chance of his being cooperative. In the present simulation, we will concentrate on situations where the trustee does not have any specific information about her current counter-

[2] The payoffs in Figure 1 reflect the *all-out* utilities governing the agents' choices. The *material* structure of a trust game (Berg et al. 1995) is usually described with the left side of Figure 1 having *(no-trust, abuse)* as unique Nash equilibrium.

[3] Cooperative preferences of the trustee can be motivated by several factors as for example social norms, sanctions, reputation, or anticipation of future interactions (Bicchieri, Xiao and Muldoon 2011). For our current purposes, the exact reasons or motivations underlying the trustee's behavior are irrelevant. We only need two types of trustees with different dominant strategies.

part, other than a general belief about how agents in society behave. We thus study the evolution and emergence of *generalized* trust.

Figure 1: Trust Game Tree



Trust Game between a trustor A and a trustee B. The utilities satisfy $A_{tr} > A_{no} > A_{def}$ for the trustor A and $B^1_{tr} < B^1_{def}$ *but* $B^2_{tr} > B^2_{def}$ for the trustee. Trustees in the left branch will exploit trust while trustees in the right branch have a dominant strategy of cooperation.

Given that trust is a rational belief held by individuals, it is as yet unclear why there are huge differences in the distribution of trust within and between nations. Regarding the determinants of trust, research has identified a variety of institutional, economic, social, and personal factors related to high levels of trust.[4] This literature widely demonstrates that institutional factors in particular can explain significant parts of the differences in trust between nation states. Next to these institutional explanations, cultural variables such as the level of education or the predominant religious conviction have also been linked to differences in trust on the macro level.

However, some observations about trust cannot be accounted for by macroscopic variables. Even when cultural and institutional factors are similar, for instance within a given country, we still find large differences in the individual level of trust. Trust can, for example, differ between neighborhoods within a city even though these do not differ in any of the aforementioned institutional and cultural factors. It is on such explanations of trust at a local level that we focus in this article.

A further important result from the empirical literature on trust is that generalized trust, once established, is stable (Bjørnskov 2007, 3). A high level of trust is maintained, even when some of the relevant macroscopic factors vary. In other words: trust stabilizes trust. While Bjørnskov (2007, 1) emphasizes

[4] See Bjørnskov (2007), Nannestad (2008), and Welch et al. (2005) for an overview over the current literature on determinants of trust.

"the importance of taking endogeneity seriously," a theoretical explanation for this pattern is as yet still missing. Or, as Nannestad (2008) puts it: "The question of trust is a huge puzzle that is not even near solution." Obviously there must be a mechanism that leads to a lock-in on high or low levels of trust. Furthermore, being independent from the relevant institutional, cultural, and personal factors, this mechanism must be situated at a micro level.

In this paper, we present a simulation model that tracks the emergence and stability of trust through micro-level behavior. The model is based on a variety of meso- and micro-level determinants identified as relevant in the current research literature. First, being embedded in stable social surroundings should help the agents correctly assess the prospect of trusting. Social context stability should thus have a relevant influence on the long term development of trust (Coleman 1994). In particular, we expect high mobility to be detrimental to the emergence of trust. Second, the level of trust initially existing within a society should have an impact on that society's level of trust, even in the long run. This initial trust level is classically understood as a form of cultural heritage, passed on from generation to generation by means of socialization. Differences in such cultural heritage might arguably explain why some populations display a high and stable level of trust and others do not (Putnam et al. 1994). Notably, such explanations based on a joint cultural heritage might work on different levels, from local neighborhoods up to nation states. And, to state the all too obvious, we would, of course, expect a close relation between the average trustworthiness, i.e. the share of trust abusers in the population, and the overall trust level.

3. Description of the Model

The present model's core is formed by trust games played between a trustor and a trustee.[5] Each pair of agents will play only a single trust game, before both parties involved move forward to engage with new partners. Over time, each agent can assume both roles, sometimes acting as trustor and sometimes as trustee. Crucially, in both of these roles, the agent can gather new information about the prospects of trust. We describe the behavior of agents in three steps, focusing on their behavior as trustees, their actions as trustors and, finally, the learning rule they use to incorporate new information.

3.1 Agents as Trustees

The easiest of these components is the agent's behavior as trustee. Within any single trust game, the trustee has no *material* reason to act cooperatively. To

[5] The model is available at <https://www.openabm.org/model/6002/version/1/view> (Accessed January 29, 2018).

the contrary, the material payoff of exploiting trust is assumed to be strictly higher than that of rewarding trust. However, there are various reasons that might lead a trustee to act trustworthily. For one, Bicchieri et al. (2011) have argued that the behavior as trustee, unlike the trustor's actions, may be guided by a social norm prescribing a reciprocation of trust. Other reasons that could motivate trustees to act trustworthily include altruistic traits or a fear of legal prosecution. All these factors have in common that they evolve on a much larger time scale than the relevant beliefs for deciding whether or not to trust. For this simulation, we make the idealizing assumption that each individual's behavior as trustee is constant over time. Each agent can either be trustworthy, always playing cooperatively, or untrustworthy, always exploiting trust. To be clear: different trustees can follow different strategies; we simply assume that each individual agent always sticks to the same strategy throughout a simulation run. Thus, within the strategy tree depicted in Figure 1, the right and left side of the tree correspond to trustworthy and untrustworthy trustees.

3.2 Agents as Trustors

Agents' behavior as trustors is guided by the rational choice approach outlined above. A trustor is willing to place trust if she expects this on average to be advantageous for her. Her behavior thus crucially depends upon her expectations towards trustees. We model this belief by a parameter, *trust expectation*, ranging from 0 to 1. The higher this value, the higher is the agent's degree of belief that others are, in general, trustworthy. This parameter may thus be interpreted as a subjective belief in the trustworthiness of others. In extreme cases, a trust expectation of 0 expresses the belief that trustees would defect for sure. A trust expectation of 1, conversely, stands for the belief that trustors are univocally trustworthy. So how does this graded belief feed into trustees' actions? Following our rational choice approach, agents place trust in others if the expected return of doing so is higher than the expected return of not doing so. The Harsanyi transformation of the trust game (see Figure 1) reduces this expected utility calculation to the simple question of how likely it is that the opponent is trustworthy. The higher the chance of being paired with a trustworthy agent, the higher the expected payoff of placing trust. Thus, for a boundedly rational agent, there will be some threshold of trust expectation from which on placing trust becomes the dominant strategy. For this simulation, we set the threshold to 0.5. That is, our agents apply the following rule:

Decision Rule (DR):
Place trust if trust expectation ≥ 0.5*, else do not play.* (DR)

Hence, an agent will agree to place trust if her trust expectation exceeds the above threshold. In all other cases she will refuse to accept the trustor's role.

3.3 Agents and Learning

The third part of our model represents the learning mechanism employed by the agents. Each agent gradually updates her beliefs about the expected trustworthiness of others. With every trust game played, an agent collects a new piece of information I about whether or not it pays off to place trust. Of course, that information will impact the agent's trust expectation and thus her behavior in future interaction. We will elaborate shortly about the types of information an agent can gain. First, let us describe how any piece of information is incorporated into the agent's trust expectation. For our analysis, we assume that freshly incoming information is binary: A new piece of evidence indicates that others are, in general, trustworthy ($I = 1$) or not ($I = 0$).

This information will then be incorporated into the agent's prior beliefs by way of a weighted average, in line with the paradigm of Bayesian sensor integration (Körding and Wolpert 2004). Naturally, agents may differ in how much weight they are willing to attribute to new information. This difference might be caused by personal preferences as well as characteristics of the environment. An agent who is uncertain about the value of information already held, for instance, may place much weight on new evidence. The same is true vice versa. We represent the weight attributed to new information by a sensitivity parameter s. The updated trust expectation is then given by the formula (I):

$$trust\ expectation_{new} = (1 - s) * trust\ expectation_{old} + s * I. \quad \text{(I)}$$

Thus the sensitivity parameter s describes how much weight is attributed to the newly received piece of evidence.

Let us now move from incorporating new information to specifying which types of information agents can acquire. There are two different ways in which agents can learn about the world. First, when acting as trustor, an agent has direct access to a new piece of evidence I about the behavior of trustees: If the current trustee cooperates, the trustor receives a positive feedback ($I = 1$). A defecting trustee, on the other hand, triggers a negative feedback ($I = 0$). However, there is also an indirect or social way of obtaining information about trustworthiness: When assuming the role of a trustee, an agent can observe whether or not the corresponding trustor places trust in her. Taking that trustor to be a rational agent playing her best strategy, this conveys some indirect or *social* clue about the prospects of trust. A trustor will place trust only if she believes that this is a rational thing to do. That is, if she has experienced trustees as predominantly trustworthy. In the baseline model, we treat this indirect way of learning on par with the direct information collected as a trustor. Thus the possible observations are again $I = 1$ if the trustor is willing to place trust and $I = 0$ if she refuses to do so. Later, we will inquire further into the subtle relationship between these two types of information, direct and social.

Finally, to conclude this section, we want to emphasize that there is no connection between an agent's behavior as trustor and trustee. Neither the decision

rule for when to place trust nor the agent's updating rule depend in any way upon that agent's behavior as a trustee. And, since the latter behavior is constant throughout time, it is obviously not influenced by her actions as trustor.

3.4 The Model

Within the present model, 1,500 agents move and interact on a two-dimensional grid, thereby gradually forming beliefs about the level of trustworthiness in the surrounding society. In order to prevent small-world effects, agents repeatedly interacting with the same partners, we work with a relatively large grid of 51 x 51 patches, populated with 1,500 randomly distributed agents. To increase homogeneity, agents crossing the right edge of the grid reappear at the left edge and vice versa – the same holds for the top and bottom edges. Each simulation run lasts for 1,000 rounds. A round consists of a first phase in which the agents interact with each other, followed by a second phase of moving around. In the interaction phase, agents randomly pick a partner who is not yet engaged in any trust game from their immediate vicinity, their von Neumann neighborhood. If no such potential partner is available, the agent stays unpartnered and does not engage in any trust game for that round. Every agent can thus be part of at most one pair at a time, acting as either trustor or trustee. After all pairs have played, all agents, including the non-partnered, move a fixed distance in a random direction. We will refer to this distance as *mobility*. Each spot can be occupied only by one agent at a time. Agents attempting to move to an already occupied field, repeat the moving routine until they find a free spot. After 1,000 rounds, the simulation stops and the final measures are extracted.[6]

Within the simulation, we systematically varied four input parameters: Three of these are the share of trustworthy agents among the population (between 40 and 70%), agents' sensitivity towards newly acquired information (between 0.03 and 0.1) and the agents' mobility, the distance each agent moves every round (between 1 and 20). Within each simulation run, these values are held constant. That is, every agent attributes the same weight to new information received and has the same moving speed. Moreover agents never change their behavior as trustees.

The fourth input parameter represents the initial trust expectation at the start of the simulation. Here, the input value describes the mean initial trust expectation of all agents (between 0.4 and 0.8). The individual agent's trust expectation is then drawn from a normal distribution around this value with a standard deviation of 0.2. We ran a total of 46,080 simulations, two runs for each com-

[6] There are certain stable configurations that the simulation does not leave once reached. In such cases we stopped the simulation early.

bination of the four input parameters. Unless noted otherwise, all data presented in the following sections is taken from these simulations.

4. Results

In this simulation we are primarily interested in the emergence and stability of generalized trust. The question to be answered here is how the level of generalized trust depends on a range of individual and societal parameters. To begin with, we study the influence of various factors identified as relevant in the current literature: Agents' social and geographical mobility, their initial trust endowment, and later the overall amount of trustworthiness present within a society. We complement our results by studying how the agents' subjective perceptions of their surroundings, whether society is stable and their peers are well informed, impact the stability of trust. Finally, we inquire into the robustness of trust towards momentary shocks, external events that diminish the agents' propensity to put trust in others.

4.1 Methodology and First Results

We are mainly interested in two issues. We want to understand the local dynamics of trust over a limited number of rounds and, second, we want to grasp an entire system's limit behavior. There are two main measures we use to address these questions. The first is the final level of trust, i.e. the share of agents willing to trust others, at the end of a simulation run. Following decision rule (DR), these are exactly the agents with a trust expectation of at least 0.5. The measure *Level of trust*, is thus defined as

$$Level\ of\ trust = \frac{number\ of\ agents\ with\ trust\ expectation \geq 0.5}{number\ of\ agents}.$$

An immediate finding is that, in the long run, the model always converges towards this measure's extreme values of 0 or 1. Eventually, either all agents become trusting or all agents become distrusting. Prima facie, this is not implausible. All agents are interested in the *same* question: whether the actual share of trustworthy agents is high enough to justify placing trust in unknown others. One could argue therefore, that all agents should arrive at the same results. However, the high degree of uniformity in convergence might come as a surprise. Despite their different learning histories, virtually all agents arrive at the same attitude. We attribute this partially to the indirect part of our learning mechanism, guided by social information. When acting as trustee, agents observe the corresponding trustor's behavior. From this, they infer the trustor's informational state and update their beliefs accordingly. Once *Level of trust* is sufficiently close to either 0 or 1, almost all trustors behave in the same way and the indirect information received by trustees is so uniformly negative (or

positive) that it drives the state of society further towards that extreme. For extreme values of the *Level of trust*, the share of trusting agents in a society thus becomes self-enforcing. The two states of universal trust and universal distrust can, therefore, be seen as stable behavioral equilibria of the iterated learning process about trust within a society. We will call simulation runs that arrived at these two stable equilibria *trusting* and *distrusting* respectively. However, we are not only interested in which equilibria exist, but also in how and when the simulation converges to either equilibrium. That is, we want to know how often the different simulations converge towards a state of universal trust and how that reflects the model's input parameters. Our second output measure, *Share trusting*, is the share of simulations that converge towards the trusting equilibrium. Given a set of simulations *S*, this output measure is defined by

$$Share\ trusting = \frac{number\ of\ trusting\ simulations\ in\ S}{number\ of\ simulations\ in\ S}.$$

It is exactly this study of equilibrium selection and equilibrium convergence with which computational models are most helpful. While classic game theoretic analysis has little to say about the dynamic processes leading towards the different equilibria, this simulation does not only offer a reproduction of the game theoretical equilibrium results, but also allows for additional insights into the way these equilibria arise.

4.2 Trust and Mobility

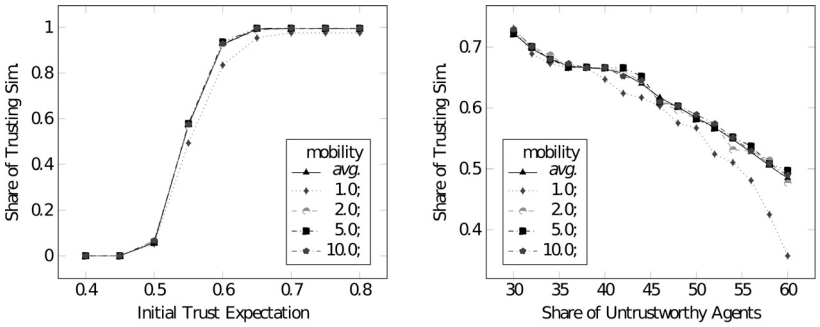
A first, key result is that agents' mobility has a strong and unexpected influence on the emergence and stability of trust. Within the simulation, mobility denotes the number of steps an agent moves after each round of interactions. This distance is the same for all agents within each simulation run. Low levels of mobility thus represent geographically and socially static societies, while dynamic societies correspond to higher mobility levels. As will become clearer later, the influence of mobility is most intricate. This parameter acts directly, through fostering or impeding the emergence of generalized trust as well as indirectly, moderating the role of other parameters and events. The role of mobility will be highlighted throughout the different sections of this paper.

First, we look at the *direct* influence of mobility, specifically at the relationship between this parameter and the share of simulation runs converging towards the trusting equilibrium.

As displayed in Figure 2, mobility is positively correlated with the level of trust. More specifically, a low mobility of 1 is detrimental to the emergence of

trust, while higher levels of mobility do not have any traceable impact.[7] Notably, all data presented here is based on the whole set of 46,080 simulations. Each data point in Figure 2 thus is an average over a set of simulation runs for the different values of the remaining parameters. The largest mobility influence occurs at moderate levels of the initial trust endowment, between 5 and 6. For higher or lower values of that parameter, the simulation is already too heavily bent towards universal trust or distrust.[8]

Figure 2: Effects of Trust and Trustworthiness



The observation that mobility favors the emergence of trust stands in contrast to a variety of results from current literature. In recent theories on social capital, the mobility factor is even believed to have a negative impact on the general level of social trust. There, it is argued that social norms and general trust tend to be stronger in smaller contexts (Putnam 1995). Mobility is sometimes even identified as the central characteristic of modern society responsible for the decline of social trust. Thus, it is worth considering the mechanisms driving mobility's influence in the current simulation. To do so, we proceed in two steps. First, we show that low mobility is correlated with a clustering of trusting and distrusting agents. As a second step we then argue how such clustering can favor the emergence of distrust.

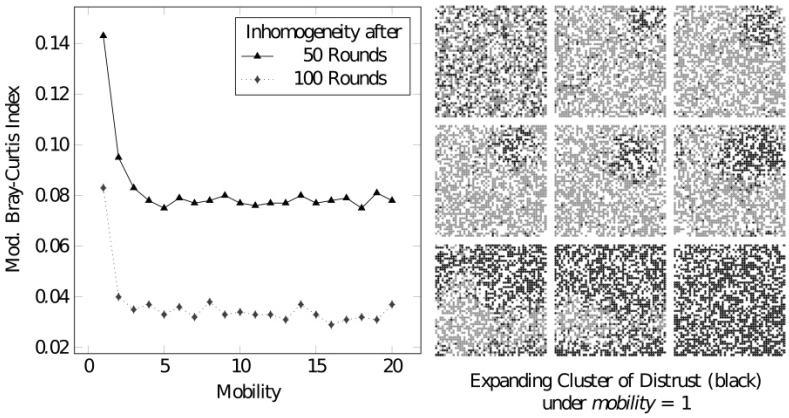
4.2.1 Mobility and Clustering

We claim that the observed difference between a mobility of one and higher levels of mobility can be traced back to a local clustering effect.

[7] Within larger grids, similar but weaker effects can be observed for higher mobility levels of 2 or 3. Thus, we take this phenomenon to be caused by the interplay of field size and mobility rather than the latter factor alone.

[8] Similar observations can be made about the share of trustworthy agents present in the simulation. Generally speaking, extreme values of these two parameters determine the outcome of the simulation and block the influence of other parameters.

Figure 3: Clustering and Mobility



With a mobility of one, local clusters of trust and distrust occur, as illustrated in the right half of Figure 3, while higher levels of mobility give rise to a more isotropic distribution of trust and distrust. We continue the argument in two steps. First we show that the emergence of trust is inversely correlated to the existence of local clusters before arguing how a local clustering can contribute to the emergence of distrust. To show that clusters primarily occur at a mobility of 1 and, much less so, at a mobility of 2, we calculate the index of dissimilarity between trusting and distrusting agents, measuring how unevenly these two types are distributed across the entire field. For our current purpose, we measure dissimilarity with the modified Bray-Curtis Index of Similarity,[9] based upon a subdivision of the field in 3 x 3 square districts of equal size. The left half of Figure 3 displays the index of dissimilarity taken after 50 and 100 simulation rounds for different values of mobility. As claimed before, low mobility contributes to a high degree of clustering, especially at the outset of the simulation. Remarkably, this clustering is completely endogenous and not reducible to any special variable or mechanism within the model. None of the mechanisms included directly fosters a clustering between trusting and distrusting agents.

[9] Let M be a map divided into a set I of different sectors and let p and q be two populations on that map. For each sector $i \in I$ let p_i and q_i be the number of p and q agents respectively living in that sector. Then the modified Bray-Curtis index of similarity between p and q is given by $\frac{1}{2}\sum_{i\in I}\left|\frac{p_i}{\sum_{j\in I} p_j} - \frac{q_i}{\sum_{j\in I} q_j}\right|$. Thus, the modified Bray-Curtis index measures the classical Bray-Curtis dissimilarity (Bray and Curtis 1957) between the local density functions $\frac{p_i}{\sum_{j\in I} p_j}$ and $\frac{q_i}{\sum_{j\in I} q_j}$.

4.2.2 Towards an Explanation of Mobility

In this section we analyze how a local clustering could favor the emergence of distrust. More precisely, we show that clusters of distrust gradually spread out, thereby infecting the trusting regions around them. See Figure 3 for an illustration. The main reason for this is that distrust is more stable than trust, for which we identify two related explanations.

First, we argue that the respective beliefs will be more extreme in distrusting clusters than in their trusting counterparts. When trusting and distrusting clusters clash, the distrusting agents will thus be more resilient in their beliefs, making it more likely for them to convert others than being converted themselves. Once a group of agents has converged into a general state of distrust, agents refuse to place any further trust. The only incoming information available to such agents is the social information collected as trustees. Within a *distrusting* cluster this information is uniformly negative, indicating that nobody is willing to place any trust. Hence, all information received is negative and the general trust expectation within a cluster will, therefore, gradually decline towards 0, the absolute minimum. For the same reasons, the social information available within a *trusting* cluster is uniformly positive as all agents are willing to place trust. However, unlike in distrusting clusters this is not the only source of information. Rather, trusting agents also continue to collect new direct experience through trusting others. Hence, the uniformly positive social information is always accompanied by pieces of direct experience. But the latter will sometimes be positive and sometimes negative, depending on the partner's trustee type. In particular, receiving occasional pieces of negative information, agents within the trusting cluster will never converge to the maximal trust expectation of 1. The average trust expectation in trusting clusters will, therefore, be less extreme than in their distrusting counterparts. This asymmetry affects the interplay of trusting and distrusting agents located at the border areas between the respective clusters. Being less entrenched in their beliefs, trusting agents are less resilient to attitude changes than their distrusting peers. It is more likely that a trusting agent becomes distrusting than vice versa, and thus the distrusting cluster gradually grows as agents update.

A second reason for the stability of distrust is related to agents' learning speed. We will argue that trusting agents update their trust expectation more often than distrusting agents. Since every change in trust attitude is preceded by some informational change, it is therefore more likely in a given time span that a trusting agent will change her mind than a distrusting agent. In general, the newly collected information of agents will be mixed, containing negative and positive pieces of evidence. This holds especially true at the borders between trusting and distrusting clusters, where both direct and indirect information can be positive or negative, depending on whether the corresponding partner is trustworthy (respectively trusting) or not. It is only through such information

that agents can change their trust attitude. However, note that trusting agents collect double as much information as distrusting agents in the same time interval. The former collect new information in the roles of trustors and trustees, while the latter use only the trustee role to update their beliefs.[10] Hence, the chance of a trusting agent switching her state in any given time interval is higher than the chance of a distrusting agent doing so, simply because the former updates her trust expectation twice as often. Thus, within any fixed time span, we should expect more trusting agents to become distrusting than vice versa, causing the distrusting population to grow gradually.

4.3 Initial Trust Endowment and the Share of Defecting Agents

In the next step of analyses, we consider two parameters related to the actual trustworthiness of a society: the number of defecting agents and the initial trust endowment at the beginning of a simulation run. The latter is an individual parameter that reflects an agent's prior experience or her socialization. Hence, different agents will start with different initial trust expectations. In each simulation run, the agent's initial trust endowment will be drawn from a normally distributed random variable with a standard variation of 0.2. The mean of this distribution, that is the average initial trust endowment, is one of the input parameters of the simulation. The left side of Figure 2 shows the influence of mean initial trust expectation on the share of simulations converging towards the trusting state. Naturally, an all too low initial trust endowment should hamper the emergence of trust.

If agents are unwilling to place trust even once, they can never collect any new information about the level of trustworthiness present, at least not by direct learning. A fortiori, if most agents start with a low trust expectation, the indirect information collected will also be overwhelmingly negative, thus reinforcing agents' negative attitude. This reasoning helps to explain the left half of Figure 2: Surprisingly, even more is true. For sufficiently high levels of initial trust endowment, agents set out to collect new information round after round. This new information gradually overwrites whatever initial expectation agents start with. Yet, despite collecting new evidence for 1,000 rounds, the initial trust endowment still has a strong influence on the emergence of trust. If the initial trust endowment is high enough, almost all simulation runs converge towards the trusting equilibrium, independent of other parameters' values (see Figure 2). This result is especially striking when contrasted with influence from the actual share of trustworthy agents. As could be expected, there is a strong

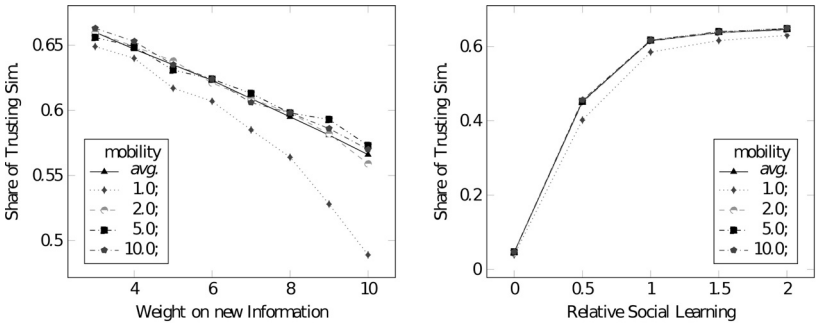
[10] Agents have an equal chance of being the first or second party in a trust game. It is only *inside* such a trust game that distrusting agents prefer to play *no trust*, preventing them from learning the opponent's attitude. Hence, being distrusting does not increase the chance of being picked as trustee.

negative correlation between the emergence of universal trust and the share of untrustworthy agents, see Figure 2. This influence is, however, much weaker than that of the initial trust endowment. The subjective perception of agents at the beginning of the simulation has a higher influence on the long term level of trust than the actual number of untrustworthy agents.

4.4 Individual Factors: Self Confidence, Group Expertise and Social Learning

We now turn our attention towards the agents' personal situation. The general level of trust is not only guided by which information agents receive, but also how this information is incorporated into their trust expectations. This, in turn depends on agents' attitudes towards their surroundings. If agents perceive surroundings to be insecure or quickly changing, they will put much weight on their most recent encounters. Additionally, if they have little faith in the knowledge or competence of others, they might put less weight on their social information and more on direct evidence. Conversely, if agents perceive their surroundings to be highly stable, they may put considerably less weight on new encounters, treating them as anecdotal cases that should not obstruct the wider perspective. If they also gain confidence in their peers' judgments, they might place much emphasis on social information so as to profit from the experience of others. Both factors, the relationship between new and old information and between direct and social information, have a bearing on the long term dynamics of trust. In other words, the level of trust is influenced by how stable and predictable agents perceive their surroundings to be.

Figure 4: Perception and Social Learning



Our agents acquire their individual levels of trust through an iterated adaptive learning process. In this process, each piece of incoming information is incorporated into agents' trust expectations through a weighted average. The weight attached to the latest piece of evidence, factor *s* in formula (I), is the agents' sensitivity towards new information. The higher this factor, the more evidential

weight an agent puts on her most recent encounters. This factor reflects agents' attitudes towards their surroundings. In insecure or quickly changing environments, agents will rely less on their own, previous assessment and factor in new information strongly. They will exhibit a high sensitivity to new information. Conversely, in static surroundings agents may to a great extent depend on their previous knowledge, paying less attention to incoming information. Experimental results in social psychology suggest that agents' sensitivity to new information is between 3 and 10%, depending on the characteristics of the situation (Bereby-Meyer and Erev 1998).

Within each simulation run, all agents share the same sensitivity to new information. That is, all agents agree in their perception of whether the surrounding society is changing slowly or rapidly. However, agents' sensitivity varies between different simulation runs. In our simulation, we find a negative relation between social sensitivity and the probability of a simulation running towards the trusting equilibrium. Agents' perception of a rapidly changing society is already detrimental to the emergence of trust. Notably, this influence is caused by agents' learning rule alone. There is no explicit mechanism making agents more cautious in dynamic surroundings. Furthermore, we find that the negative impact of high sensitivity to new information is especially strong when paired with a low mobility. While agents' mobility has little influence when sensitivity is low, a mobility of one combined with a high level of sensitivity radically reduces the long term level of trust.

The reason for social sensitivity's influence can be traced back to the asymmetry between trusting and distrusting agents. Lower sensitivity makes agents less dependent on more recent information and more focused on the wider implications of the information collected. Arguably, this focus on the bigger picture fosters the creation and stability of trust, as long as there are, in fact, enough trustworthy agents. No matter how high the trust expectation might be, an unfortunate series of trust-abusing encounters can always move agents into a state of distrust from which it is difficult to escape. Through a combination of chance and bad luck, every trusting agent can encounter a sequence of negative information long enough to thwart her trust. The shorter the length of such a 'behavior flipping' sequence, the more likely it is to occur and hence the easier it is for an agent to lose her trust. But the higher the social sensitivity, that is the more weight an agent puts on her most recent encounters, the less consecutive negative encounters are needed to destroy the agent's trust. Hence, an increased social sensitivity raises the chance of agents' encountering a trust thwarting sequence of negative encounters.

In the basic simulation, direct and social information are treated on par. Agents attribute the same weight to each piece of evidence, be it collected through direct, first order experience or through social learning from the behavior of another trustor. Of course, this need not be the case. If trustors are skeptical about the motivations or competence of others, they may attribute less

importance to social information. Conversely, if agents take their peers to be well informed, they might rely heavily on social evidence. Since each piece of social information reflects an entire history of evidence collected by the respective trustor, this type of evidence could even be seen as more valuable than a single piece of direct information. In short, the weights attributed to direct and social information reflects, among other things, agents' perception of each other. The more competent others are taken to be, the higher is the weight attributed to social information.

In a second set of experiments, we are interested in how the balance between direct and social information impacts the stability of trust. For this, the original simulation is augmented with an additional factor controlling the relationship between direct and social learning. The factor *Relative social learning* denotes how much weight is attached to social information, relative to direct information. Formally, this means[11]

$$Relative\ social\ learning = \frac{weight\ on\ incoming\ social\ information}{weight\ on\ incoming\ direct\ information}.$$

Thus, a relative weight of 0 means that social information is not taken into account at all, while a relative weight of 1 indicates that direct and social information are treated on par. Within each simulation run, all agents share the same value of *Relative social learning*. That is, all agents agree in their subjective attitudes towards the trustworthiness of others. However, we vary this factor between different simulation runs.

We find that the relative role of direct and social information does influence the long term level of trust. The higher the emphasis on social information, the more likely a society is to establish and maintain long term trust. In other words, agents' perception of their peers impacts on the level of trust. When others are seen as knowledgeable and reliable, a society is more likely to develop and maintain generalized trust. As in the case of social sensitivity, this observation is driven by the learning rule alone.

The reason for this result is that social information always favors the majority opinion. If most agents are willing to trust others, social information will, on average, be positive, pushing agents towards trust. Likewise, if most agents are disinclined to trust others, the average indirect information will be negative, discouraging receivers from trusting others. Taken together, these considerations suggest that social information always reinforces the current majority opinion. Moreover, the higher the emphasis on social information is, the stronger will be the influence of the majority opinion. As argued above, simulations starting with a majority of distrusting agents univocally converge to distrust already, so the majoritarian pressure of increased social attention cannot

[11] Technically, we have replaced the original updating formula (I) with $trust\ expectation_{new} = (1 - s * f) * trust\ expectation_{old} + s * f * I$. Here, f is the relative social learning factor.

be reflected in the results. Rather, social learning is able to have a measurable effect only if the majority of agents start in a trusting state, in which case, social learning favors the emergence of trust.

Remarkably, without any social learning, i.e. at a relative weight of 0, all simulations univocally converge to a state of universal distrust. If the simulation runs long enough, all agents will, with a probability of 1, eventually have a sequence of negative encounters long enough to move them into distrust. But once an agent has lost her propensity to trust, she cannot regain this propensity as only indirect learning would allow her to do so. Thus, social learning is the central driving factor for the emergence and stability of trust.

The way agents perceive their surroundings impacts on the emergence of trust. When peers are seen as knowledgeable and the situation as stable, trust is likely to emerge. When agents are wary about the competence of their fellow trustors or exposed to changing surroundings, this significantly reduces the prospects of trust. However, neither of these two factors impacts on the quantitative interplay between the remaining parameters as presented in the previous sections. All qualitative relationships presented hold, irrespective of agents' perception of their surroundings.[12]

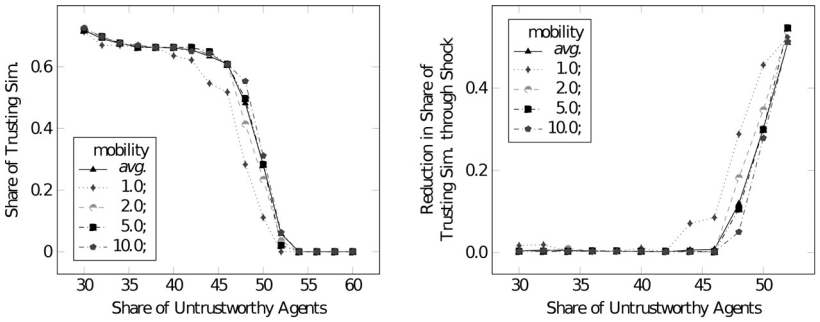
4.5 The Robustness of Trust

Finally, we turn our attention to the robustness of trust. How resilient is universal trust to informational shocks, short term intensive events that significantly change agents' perception of their surroundings? Sometimes agents base their expectations about trust on more than their own direct and social information. When a prominent event, say a group of fraudsters exploiting innocent citizens, makes its round through the media, this will impact on many people's attitudes towards trust. If such an event receives enough coverage, it can easily have a strong enough influence to completely reverse some agents' trust attitudes. Of course, the same holds for spreading rumors or running cleverly designed media campaigns. In line with the literature on informational feedback phenomena

[12] To test the validity of our findings, we ran two further samples of 46,080 simulations with the same starting parameters as the original simulations, but incorporating two slight variations. In the first variation, we replaced the assumption that there are no social or geographical limitations to agents' mobility by introducing spatial restrictions. While ensuring that every position could, in principle, be reached by every agent, spatial restrictions create secluded regions with bottle-neck accesses that make it hard to enter or leave these areas. The second variation increased the field size by a factor of 16, adjusting the number of agents accordingly. Unsurprisingly, both variations display some new results. In the first variation it is no longer true that every simulation converges to universal trust or distrust. Rather, the different secluded regions each converge to their own, local homogeneous state of universal trust or distrust. In the second variation, the detrimental influence of low mobility also shows at a mobility level of two. All of the qualitative relationships between the different influence factors reported here equally show in both these variations.

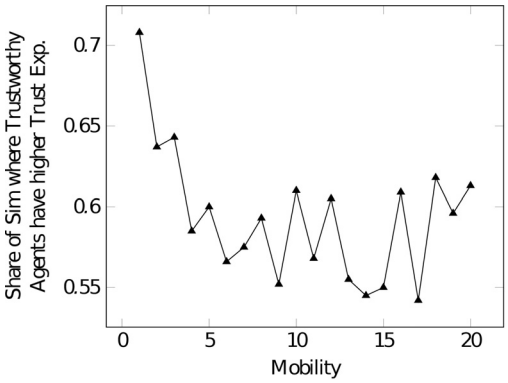
(Hansen, Hendricks and Rendsvig 2013), we will refer to all these events as informational shocks, singular events impacting most agents' beliefs about trust. Notably, informational shocks do not change the actual level of trustworthiness. They merely change agents' subjective beliefs about the value of trust.

Figure 5: The Impact of Shocks



In principle, the influence of informational shocks can go in either direction, raising or lowering the general perception of trustworthiness. For now, we focus our attention on negative shocks, short term events that thwart agents' trust expectations. After 200 rounds of simulation, we let an external shock diminish each agent's trust expectation by a random amount between 0 and 0.5 points. We then enquire whether or not such a short term intensive shock can have a lasting effect on the level of trust. That is, we ask whether or not negative shocks can convert some high trust scenario close to universal trust into a state of universal distrust. As it turns out, universal trust, once attained, is highly robust. The shocks studied in this scenario are rather severe. The maximal possible impact of 0.5 suffices to turn even the most optimistic agent into distrust. And of course, most agents will not even have such an optimistic trust expectation, as they frequently encounter untrustworthy trustees. Yet, in many cases, informational shocks do not impact on the long term behavior of a society at all. Figure 5 displays the long term behavior of simulations with shocks. As always, final measures were extracted after 1,000 simulation rounds, so long after the shock in round 200. The left side of Figure 5 displays the share of simulations that, in the long run, converges to the trusting equilibrium. The right side of Figure 5 displays by how much shocks impact the long term prospect of trust. More specifically, this graph is the difference between long term trust behavior without shocks (Figure 2) and with shocks (right hand side of this Figure). As can be seen, it takes a high number of untrustworthy agents, more than 40%, for shocks to have any long term impact.

Figure 6: Trust Level of Trustworthy and Defecting Agents



Furthermore, it turns out once again that low mobility has a negative influence on the robustness of trust. When taking the difference between the number of simulations converging towards universal trust with and without shocks, we find that mobility creates a strong impact. At a mobility of 1, this difference is much larger than for higher values. In other words, not only does low mobility impede the emergence of trust, but it also makes trust more vulnerable to short term informational shocks.

4.6 Trust and Trustworthiness, a Puzzle

We end this result section with an unexpected observation about the relation between trust and trustworthiness. Many theoretical frameworks postulate a correlation between an agent's behavior as trustor and trustee. In particular, various theoretical accounts (see e.g. Falcone and Castelfranchi 2004) claim that increased trust in others fosters one's own propensity to be trustworthy and, vice versa, being trustworthy is correlated with higher expectations towards others. No such mechanism is implemented in the current model. To the contrary, each agent's trustee type is fixed throughout a simulation run while her behavior as trustor is guided exclusively by past experience. Furthermore, the model does not contain any retaliation mechanism against trust abusers. In fact, agents do not remember about their past interaction partners, nor can they identify them in any way. Yet, the simulation displays a significant correlation between agents' trustworthiness and their propensity to trust others.

If the two roles of trustor and trustee are independent of each other, it should be as likely as not that the trust expectation of trustworthy agents is higher than that of untrustworthy agents. However, as shown in Figure 6 this does not hold true. Rather, trustworthy agents have higher levels of trust than their untrustworthy peers in far over half of the simulations. The effect is strongest at low mobility levels, with trustworthy agents outperforming abusers in up to 75% of

all simulations. We conjecture that this is caused by a local echo chamber effect. Being trustworthy and thereby producing positive feedback to one's immediate surroundings increases the likelihood of later encountering trusting agents when acting as trustee. Thus, being trustworthy can influence the content of future information, even if no agent engages in any type of personalized learning. This conjecture is supported by the observation that the effect is at its strongest when local surroundings are relatively stable and echo effects have time to pan out. This, of course, is the case when mobility is low.

5. Conclusion and Outlook

Generalized trust is a driving factor for the economic and political success of societies. Consequentially, the factors and determinants responsible for high levels of trust have, to date, been the subject of a vast body of empirical and theoretical research. Social theories, however, usually have problems in capturing the procedural character of social life. This is particularly true in the case of trust, where Bjørnskov (2007) and Nannestad (2008) remark that a theoretical explanation for the dynamics of trust is still lacking. Bjørnskov (2007, 1) even goes a step further by stressing that endogeneity plays a crucial role in understanding the driving mechanisms of trust. In this paper, we addressed this gap by means of a computer simulation.

It goes without saying that the current simulation abstracts away from various real life factors that bear on the emergence of trust. The model, for instance, captures neither that information may travel in social circles nor that there may be differences in trustworthiness between high and low mobility populations. It is exactly through ignoring such factors, that the model can shed light on various subtle factors and mechanisms relevant to the dynamics of trust. Our results on the influence of mobility, for instance, suggest that any *negative* correlation between mobility and trust, as is often reported in the literature, could be exclusively determined by differences in trustworthiness between groups of high and low mobility. We have here provided a variety of insights into the interactive, dynamic mechanisms guiding the emergence of trust. To obtain the full picture, these need to be combined with other insights on the distribution of trustworthiness or in-group dynamics.

There are three main findings that we want to emphasize: First, we observe a close link between the mobility of a society and its long term level of trust. Ceteris paribus, immobile societies are much less likely to develop and sustain high levels of trust than their more mobile counterparts. Furthermore, low mobility also increases long term vulnerability to informational shocks and other external influence. Second, agents' subjective perception also has a bearing on the emergence of trust. When agents trust the judgment of their peers and see society as stable, trust is likely to emerge. Conversely if agents per-

ceive society as quickly changing and their environment as ill-informed, we should expect much lower levels of trust. Crucially, these results occur even without any inherent bias from agents. It is an unfortunate consequence of agents' learning mechanism that creates a connection between their subjective perception of the environment and the emergence of trust. The third core insight is that the long-term trust-level of a society is determined to a large extent by its initial trust endowment. In fact, we find that agents' initial trust assessment, sometimes referred to as cultural heritage, has larger bearing on the dynamics of trust than the average level of trustworthiness itself. More generally, both widespread trust and distrust, once achieved, are resilient towards a variety of factors and events. This finding helps to shed light on the observation (see Bjørnskov 2007) that societies can sometimes display strong and lasting differences in their trust level, despite being similar in most of the relevant societal and economic determinants of trust.

On a conceptual level, the simulation highlights that a key for understanding the dynamics of trust lies in its informational structure. Concerning direct information, trust suffers from an inherent informational asymmetry. While trusting agents collect evidence with every new encounter, distrusting agents do not have access to such information. This asymmetry structurally favors distrust, as new information can always lead agents to revise their beliefs. But agents can also collect social information about trust, through observing the behavior of others. As documented in a variety of studies, such social learning tends to reinforce the majority beliefs and attitudes of groups, be they beneficial or detrimental to their members (Hansen et al. 2013). In extreme cases, social information can outweigh any other newly incoming information, thus locking in societies to states of universal trust or distrust. It is this phenomenon that ensures universal trust, once reached, to be relatively stable, both to spontaneous mutations and external shocks. But, conversely, reinforcement of distrust makes it extremely difficult to regain societal trust once lost. The interplay of informational asymmetry and social reinforcement creates complex dynamics that drive most of the results here. The interaction between these two is strongest at low mobilities, where the agents' behavior is immediately mirrored back to them, through the behavior of their surroundings. Local echo chambers are likely to form and expand. If there is any general lesson to be learned here, it is the importance of bringing agents back to participate in social life. This does not necessary mean that they will have positive experiences and learn to trust. But without participation there is not even the slightest chance of becoming trusting again.

References

Bereby-Meyer, Yoella, and Ido Erev. 1998. On learning to become a successful loser: a comparison of alternative abstractions of learning processes in the loss domain. *Journal of Mathematical Psychology* 42: 266-86.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10: 122-42.

Bicchieri, Cristina, Erte Xiao, and Ryan Muldoon. 2011. Trustworthiness is a social norm, but trusting is not. *Politics, Philosophy & Economics* 10: 170-87.

Birk, Andreas. 2001. Learning to trust. In *Trust in Cyber-societies. Integrating the Human and Artificial Perspective,* ed. Rino Falcone, Munindar Singh and Yao-Hua Tan, 133-44. Berlin: Springer.

Bjørnskov, Christian. 2007. Determinants of generalized trust: A cross-country comparison. *Public choice* 130: 1-21.

Bray, J Roger, and John T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs* 27: 325-49.

Buskens, Vincent. 2002. *Social networks and trust*. Boston: Springer.

Coleman, James S. 1994. *Foundations of social theory*. Cambridge, MA: Harvard University Press.

Falcone, Rino, and Cristiano Castelfranchi. 2004. Trust dynamics: How trust is influenced by direct experiences and by trust itself. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*-vol. 2, 740-7. IEEE Computer Society.

Hansen, Pelle G., Vincent F. Hendricks, and Rasmus K. Rendsvig. 2013. Infostorms. *Metaphilosophy* 44: 301-26.

Hooghe, Marc, and Dietlind Stolle. 2003. *Generating social capital: Civil society and institutions in comparative perspective*. New York: Palgrave Macmillan.

Hyyppä, Markku T. 2010. *Healthy ties: Social capital, population health and survival*. Dordrecht: Springer.

Knack, Stephen, and Philip Keefer. 1997. Does social capital have an economic payoff? A cross-country investigation. *The Quarterly journal of economics*: 1251-88.

Körding, Konrad P., and Daniel M. Wolpert. 2004. Bayesian integration in sensorimotor learning. *Nature* 427: 244-7.

Kornai, János, Bo Rothstein, and Susan Rose-Ackerman. 2004. *Creating social trust in post-socialist transition*. New York: Palgrave Macmillan.

Nannestad, Peter. 2008. What have we learned about generalized trust, if anything? *Annual Review of Political. Science.* 11: 413-36.

Nooteboom, Bart, Tomas Klos, and René Jorna. 2001. Adaptive trust and co-operation: An agent-based simulation approach. In *Trust in Cyber-societies. Integrating the Human and Artificial Perspective,* ed. Rino Falcone, Munindar Singh and Yao-Hua Tan, 83-109. Berlin: Springer

Ostrom, Elinor. 1990. *Governing the commons: The evolution of institutions for collective action.* Cambridge: Cambridge University Press.

Putnam, Robert D. 1995. Tuning in, tuning out: The strange disappearance of social capital in America. *PS: Political Science & Politics* 28: 664-83.

Putnam, Robert D. 2000. *Bowling alone: The collapse and revival of American community*. New York: Simon and Schuster.
Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nanetti. 1994. *Making democracy work: Civic traditions in modern Italy*. Princeton: Princeton University Press.
Sztompka, Piotr. 1999. *Trust: A sociological theory*. Cambridge: Cambridge University Press.
Torche, Florencia, and Eduardo Valenzuela. 2011. Trust and reciprocity: A theoretical distinction of the sources of social capital. *European Journal of Social Theory* 14: 181-98.
Uslaner, Eric M. 2002. *The moral foundations of trust*. Cambridge: Cambridge University Press.
Welch, Michael R., Roberto E.N. Rivera, Brian P. Conway, Jennifer Yonkoski, Paul M. Lupton, and Russell Giancola. 2005. Determinants and consequences of social trust. *Sociological inquiry* 75: 453-73.
Wilensky, Uri. 1999. *NetLogo*. <http://ccl.northwestern.edu/netlogo/> (Accessed January 11, 2018). Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.